Chapter 9
Inferences from Two Samples

Review

In Chapters 7 and 8 we introduced methods of *inferential statistics*.

In Chapter 7 we presented methods of constructing confidence interval estimates of population parameters.

In Chapter 8 we presented methods of testing claims made about population parameters.

Chapters 7 and 8 both involved methods for dealing with a sample from a single population.

Preview

The objective of this chapter is to extend the methods for estimating values of population parameters and the methods for testing hypotheses to situations involving two sets of sample data instead of just one.

The following are examples typical of those found in this chapter, which presents methods for using sample data from two populations so that inferences can be made about those populations.

Preview

- Test the claim that when college students are weighed at the beginning and end of their freshman year, the differences show a mean weight gain of 15 pounds (as in the "Freshman 15" belief).
- Test the claim that the proportion of children who contract polio is less for children given the Salk vaccine than for children given a placebo.
- Test the claim that subjects treated with Lipitor have a mean cholesterol level that is lower than the mean cholesterol level for subjects given a placebo.

Chapter 9
Inferences from Two Samples

Key Concept

In this section we present methods for (1) testing a claim made about two population proportions and (2) constructing a confidence interval estimate of the difference between the two population proportions.

This section is based on proportions, but we can use the same methods for dealing with probabilities or the decimal equivalents of percentages.

Notation for Two Proportions

For population 1, we let:

p_1 = population proportion

n_1 = size of the sample

x_1 = number of successes in the sample

$$\hat{p}_1 = \frac{x_1}{n_1}$$ (the sample proportion)

$$\hat{q}_1 = 1 - \hat{p}_1$$

The corresponding notations apply to $p_2, n_2, x_{2,}\ \hat{p}_2\ and\ \hat{q}_2$, which come from population 2.

Pooled Sample Proportion

- The pooled sample proportion is given by:

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\bar{q} = 1 - \bar{p}$$

Requirements

1. We have proportions from two independent simple random samples.
2. For each of the two samples, the number of successes is at least 5 and the number of failures is at least 5.

Test Statistic for Two Proportions

For $H_0 : p_1 = p_2$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\overline{p}\overline{q}}{n_1} + \frac{\overline{p}\overline{q}}{n_2}}}$$

where (assumed in the null hypothesis) $p_1 - p_2 = 0$

Testing Two Proportions

P-value: P-values are automatically provided by technology. If technology is not available, use Table A-2.

Critical values: Use Table A-2. (Based on the significance level α, find critical values by using the procedures introduced in Section 8-2 in the text.)

Confidence Interval Estimate of $p_1 - p_2$

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

with margin of error $E = z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

Example

Do people having different spending habits depending on the type of money they have?

89 undergraduates were randomly assigned to two groups and were given a choice of keeping the money or buying gum or mints.

The claim is that "money in large denominations is less likely to be spent relative to an equivalent amount in many smaller denominations".

Let's test the claim at the 0.05 significance level.

Example

Below are the sample data and summary statistics:

	Group 1	Group 2
	Subjects Given $1 Bill	Subjects Given 4 Quarters
Spent the money	$x_1 = 12$	$x_2 = 27$
Subjects in group	$n_1 = 46$	$n_2 = 43$

$$\hat{p}_1 = \frac{12}{46} \qquad \hat{p}_2 = \frac{27}{43}$$

$$\overline{p} = \frac{12+27}{46+43} = 0.438202$$

Example

Requirement Check:

1. The 89 subjects were randomly assigned to two groups, so we consider these independent random samples.

2. The subjects given the $1 bill include 12 who spent it and 34 who did not. The subjects given the quarters include 27 who spent it and 16 who did not. All counts are above 5, so the requirements are all met.

Example

Step 1: The claim that "money in large denominations is less likely to be spent" can be expressed as $p_1 < p_2$.

Step 2: If $p_1 < p_2$ is false, then $p_1 \geq p_2$.

Step 3: The hypotheses can be written as:

$$H_0 : p_1 = p_2$$

$$H_1 : p_1 < p_2$$

Example

Step 4: The significance level is $\alpha = 0.05$.

Step 5: We will use the normal distribution to run the test with:

$$\hat{p}_1 = \frac{12}{46} \qquad \hat{p}_2 = \frac{27}{43}$$

$$\overline{p} = \frac{12+27}{46+43} = 0.438202$$

$$\overline{q} = 1 - \overline{p} = 1 - 0.438202 = 0.561798$$

Example

Step 6: Calculate the value of the test statistic:

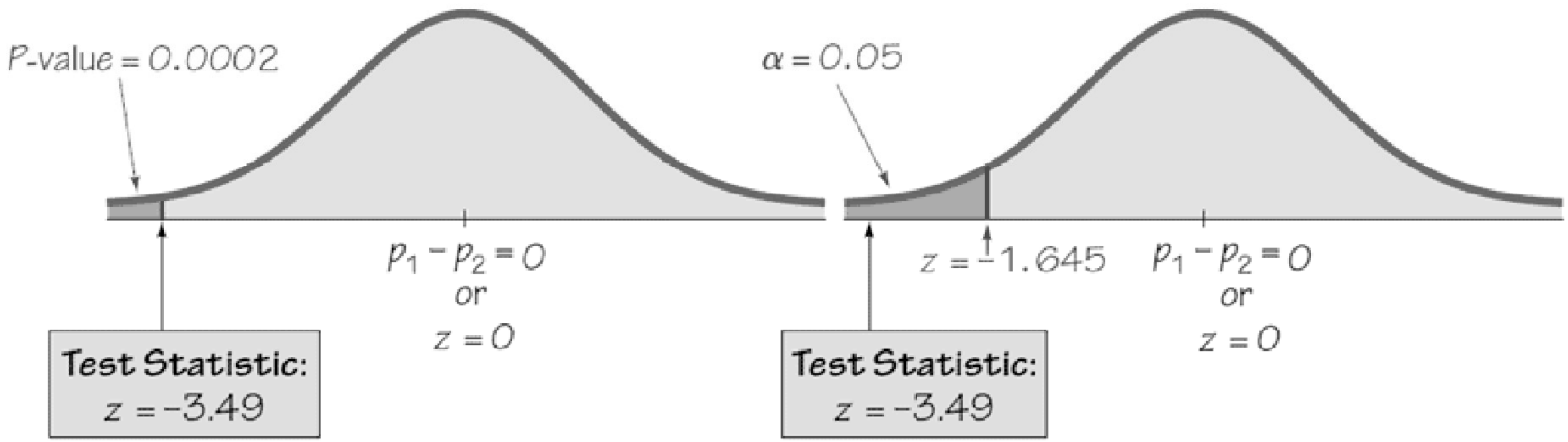
$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\overline{p}\overline{q}}{n_1} + \frac{\overline{p}\overline{q}}{n_2}}}$$

$$= \frac{\left(\frac{12}{46} - \frac{27}{43}\right) - 0}{\sqrt{\frac{(0.438202)(0.561798)}{46} + \frac{(0.438202)(0.561798)}{43}}}$$

$$= -3.49$$

Example

Step 6: This is a left-tailed test, so the *P*-value is the area to the left of the test statistic $z = -3.49$, or 0.0002. The critical value is also shown below.



(a) *P*-Value Method

(b) Critical Value Method

Example

Step 7: Because the *P*-value of 0.0002 is less than the significance level of $\alpha = 0.05$, reject the null hypothesis.

There is sufficient evidence to support the claim that people with money in large denominations are less likely to spend relative to people with money in smaller denominations.

It should be noted that the subjects were all undergraduates and care should be taken before generalizing the results to the general population.

Example

We can also construct a confidence interval to estimate the difference between the population proportions.

Caution: The confidence interval uses standard deviations based on estimated values of the population proportions, and consequently, a conclusion based on a confidence interval might be different from a conclusion based on a hypothesis test.

Example

Construct a 90% confidence interval estimate of the difference between the two population proportions.

What does the result suggest about our claim about people spending large denominations relative to spending small denominations?

Example

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$= 1.645\sqrt{\frac{\left(\frac{12}{46}\right)\left(\frac{34}{46}\right)}{46} + \frac{\left(\frac{27}{43}\right)\left(\frac{16}{43}\right)}{43}}$$

$$= 0.161387$$

Example

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

$$-0.367037 - 0.161387 < (p_1 - p_2) < -0.367037 + 0.161387$$

$$-0.528 < (p_1 - p_2) < -0.206$$

The confidence interval limits do not include 0, implying that there is a significant difference between the two proportions.

There does appear to be sufficient evidence to support the claim that "money in large denominations is less likely to be spent relative to an equivalent amount in many smaller denominations."

Chapter 9
Inferences from Two Samples

Key Concept

This section presents methods for using sample data from two independent samples to test hypotheses made about two population means or to construct confidence interval estimates of the difference between two population means.

Key Concept

In Part 1 we discuss situations in which the standard deviations of the two populations are unknown and are not assumed to be equal.

In Part 2 we discuss two other situations: (1) The two population standard deviations are both known; (2) the two population standard deviations are unknown but are assumed to be equal.

Because is typically unknown in real situations, most attention should be given to the methods described in Part 1.

Part 1: Independent Samples with σ_1 and σ_2 Unknown and Not Assumed Equal

Definitions

Two samples are independent if the sample values selected from one population are not related to or somehow paired or matched with the sample values from the other population.

Two samples are dependent if the sample values are *paired*. (That is, each pair of sample values consists of two measurements from the same subject (such as before/after data), or each pair of sample values consists of matched pairs (such as husband/wife data), where the matching is based on some inherent relationship.)

Example

Independent Samples:

Researchers investigated the reliability of test assessment. On group consisted of 30 students who took proctored tests. A second group consisted of 32 students who took online tests without a proctor.

The two samples are independent, because the subjects were not matched or paired in any way.

Example

Dependent Samples:

Students of the author collected data consisting of the heights (cm) of husbands and the heights (cm) of their wives. Five of those pairs are listed below. The data are dependent, because each height of each husband is matched with the height of his wife.

Height of Husband	175	180	173	176	178
Height of Wife	160	165	163	162	166

Notation

μ_1 = population mean

σ_1 = population standard deviation

n_1 = size of the first sample

$\bar{x}_1$ = sample mean

s_1 = sample standard deviation

Corresponding notations apply to population 2.

Requirements

1. σ_1 and σ_2 are unknown and no assumption is made about the equality of σ_1 and σ_2.
2. The two samples are independent.
3. Both samples are simple random samples.
4. Either or both of these conditions are satisfied:

   The two sample sizes are both large (over 30) or both samples come from populations having normal distributions.

Hypothesis Test for Two Means: Independent Samples

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

(where $\mu_1 - \mu_2$ is often assumed to be 0)

Hypothesis Test

Test Statistic for Two Means: Independent Samples

Degrees of freedom:	In this book we use this simple and conservative estimate: df = smaller of $n_1 - 1$ or $n_2 - 1$.
P-values:	Refer to Table A-3. Use the procedure summarized in Figure 8-4.
Critical values:	Refer to Table A-3.

Confidence Interval Estimate of $\mu_1 - \mu_2$ Independent Samples

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

where $E = t_{\alpha/2}\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$

df = smaller of $n_1 - 1$ or $n_2 - 1$.

Caution

Before conducting a hypothesis test, consider the context of the data, the source of the data, the sampling method, and explore the data with graphs and descriptive statistics.

Be sure to verify that the requirements are satisfied.

Example

Researchers conducted trials to investigate the effects of color on creativity.

Subjects with a red background were asked to think of creative uses for a brick; other subjects with a blue background were given the same task.

Responses were given by a panel of judges.

Researchers make the claim that "blue enhances performance on a creative task". Test the claim using a 0.01 significance level.

Example

Requirement check:

1. The values of the two population standard deviations are unknown and assumed not equal.
2. The subject groups are independent.
3. The samples are simple random samples.
4. Both sample sizes exceed 30.

The requirements are all satisfied.

Example

The data:

Creativity Scores			
Red Background	n = 35	$\bar{x} = 3.39$	s = 0.97
Blue Background	n = 36	$\bar{x} = 3.97$	s = 0.63

Example

Step 1: The claim that "blue enhances performance on a creative task" can be restated as "people with a blue background (group 2) have a higher mean creativity score than those in the group with a red background (group 1)". This can be expressed as $\mu_1 < \mu_2$.

Step 2: If the original claim is false, then $\mu_1 \geq \mu_2$.

Example

Step 3: The hypotheses can be written as:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

Step 4: The significance level is $\alpha = 0.05$.

Step 5: Because we have two independent samples and we are testing a claim about two population means, we use a t distribution.

Example

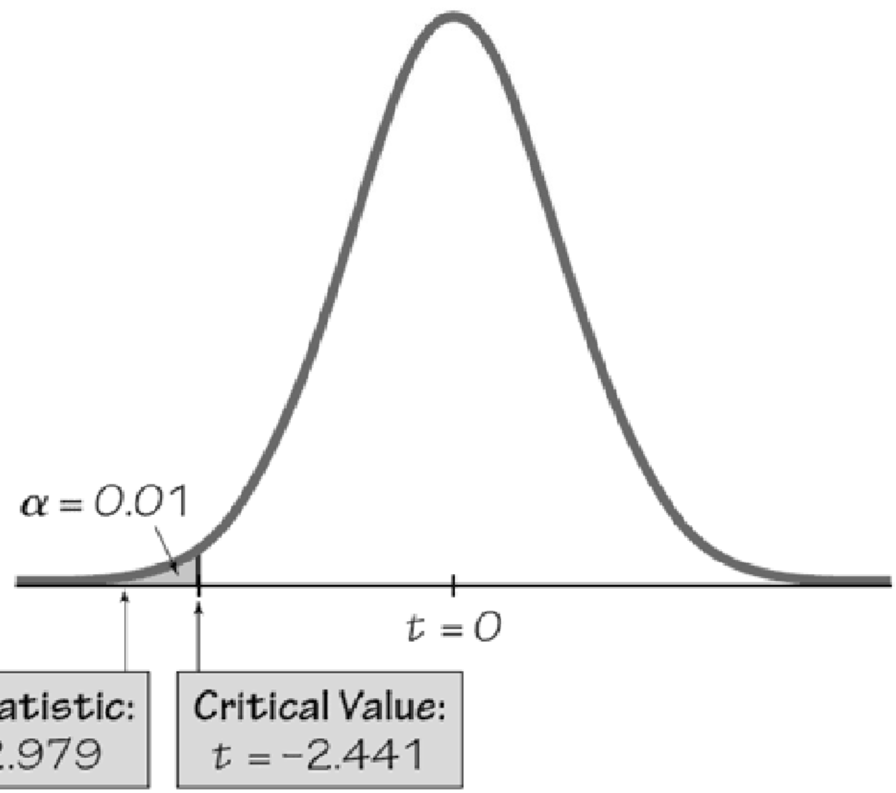
Step 6: Calculate the test statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

$$= \frac{(3.39 - 3.97) - 0}{\sqrt{\dfrac{0.97^2}{35} + \dfrac{0.63^2}{36}}} = -2.979$$

Example

Step 6: Because we are using a *t* distribution, the critical value of t = –2.441 is found from Table A-3. We use 34 degrees of freedom.



Example

Step 7: Because the test statistic does fall in the critical region, we reject the null hypothesis $\mu_1 - \mu_2$.

P-Value Method: Technology will provide a P-value, and the area to the left of the test statistic of $t = -2.979$ is 0.0021. Since this is less than the significance level of 0.01, we reject the null hypothesis.

Example

There is sufficient evidence to support the claim that the red background group has a lower mean creativity score than the blue background group.

Example

Using the data from this color creativity example, construct a 98% confidence interval estimate for the difference between the mean creativity score for those with a red background and the mean creativity score for those with a blue background.

Example

$$E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2.441\sqrt{\frac{0.97^2}{35} + \frac{0.63^2}{36}} = 0.475261$$

$$\bar{x}_1 = 3.39 \quad \text{and} \quad \bar{x}_2 = 3.97$$

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

$$-1.06 < (\mu_1 - \mu_2) < -0.10$$

Example

We are 98% confident that the limits –1.05 and –0.11 actually do contain the difference between the two population means.

Because those limits do not include 0, our interval suggests that there is a significant difference between the two means.

Chapter 9
Inferences from Two Samples

Key Concept

In this section we develop methods for testing hypotheses and constructing confidence intervals involving the mean of the differences of the values from two dependent populations.

With dependent samples, there is some relationship where each value in one sample is paired with a corresponding value in the other sample.

Key Concept

Because the hypothesis test and confidence interval use the same distribution and standard error, they are equivalent in the sense that they result in the same conclusions.

Consequently, the null hypothesis that the mean difference equals 0 can be tested by determining whether the confidence interval includes 0.

There are no exact procedures for dealing with dependent samples, but the *t* distribution serves as a reasonably good approximation, so the following methods are commonly used.

Dependent Samples

Example – Matched pairs of heights of U.S. presidents and heights of their main opponents. Since there is a relationship as a basis for matching the pairs of data, this data consists of dependent samples.

Height (cm) of President	189	173	183	180	179
Height (cm) of Main Opponent	170	185	175	180	178

Good Experimental Design

When designing an experiment or planning an observational study, using dependent samples with paired data is generally better than using two independent samples.

The advantage of using matched pairs is that we reduce extraneous variation, which could occur if each experimental unit were treated independently.

Notation for Dependent Samples

d	=	individual difference between the two values of a single matched pair
μ_d	=	mean value of the differences d for the population of all matched pairs of data
$\bar{d}$	=	mean value of the differences d for the paired sample data
s_d	=	standard deviation of the differences d for the paired sample data
n	=	number of pairs of sample data

Requirements

1. The sample data are dependent.

2. The samples are simple random samples.

3. Either or both of these conditions is satisfied: The number of pairs of sample data is large ($n > 30$) or the pairs of values have differences that are from a population having a distribution that is approximately normal.

Hypothesis Test Statistic for Matched Pairs

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

where degrees of freedom = n – 1

P-values and Critical Values

P-values: *P*-values are automatically provided by technology. If technology is not available, use Table A-3.

Critical values: Use Table A-3 with degrees of freedom df = $n - 1$

Confidence Intervals for Matched Pairs

$$\bar{d} - E < \mu_d < \bar{d} + E$$

$$E = t_{\alpha/2} \frac{s_d}{\sqrt{n}}$$

Critical values of $t_{\alpha/2}$:

Use Table A-3 with df = n – 1 degrees of freedom.

Example

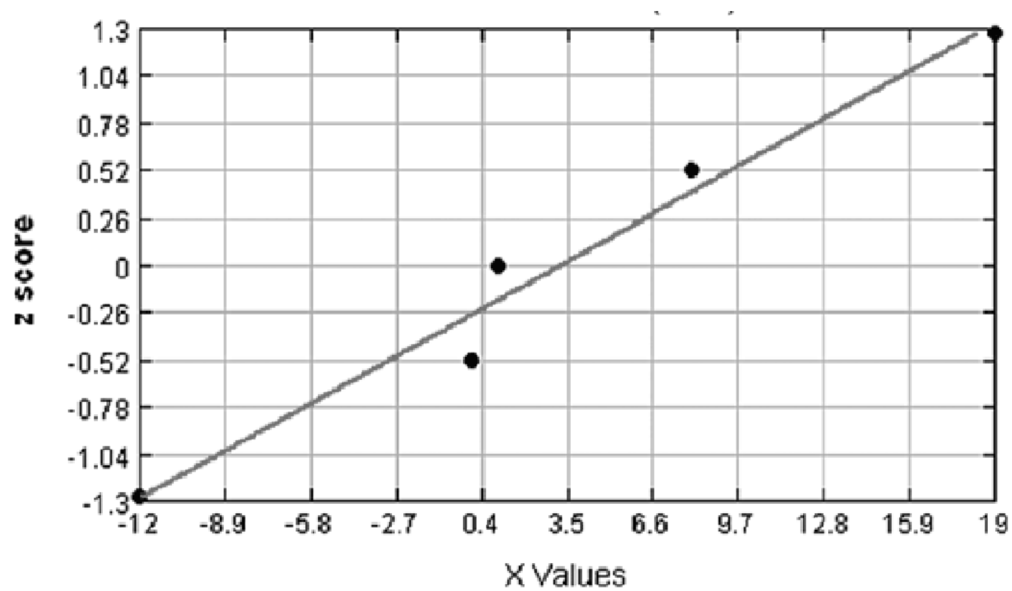
Use the sample data below with a significance level of 0.05 to test the claim that for the population of heights of presidents and their main opponents, the differences have a mean greater than 0 cm (so presidents tend to be taller than their opponents).

Height (cm) of President	189	173	183	180	179
Height (cm) of Main Opponent	170	185	175	180	178
Difference *d*	19	-12	8	0	1

Example

Requirement Check:

1. The samples are dependent because the values are paired.
2. The pairs of data are randomly selected.
3. The number of data points is 5, so normality should be checked (and it is assumed the condition is met).



Example

Step 1: The claim is that $\mu_d > 0$ cm.

Step 2: If the original claim is not true, we have $\mu_d \leq 0$ cm.

Step 3: The hypotheses can be written as:

$$H_0 : \mu_d = 0 \text{ cm}$$

$$H_0 : \mu_d > 0 \text{ cm}$$

Example

Step 4: The significance level is $\alpha = 0.05$.

Step 5: We use the Student t distribution.

The summary statistics are:

$$\bar{d} = 3.2$$

$$s = 11.4$$

Example

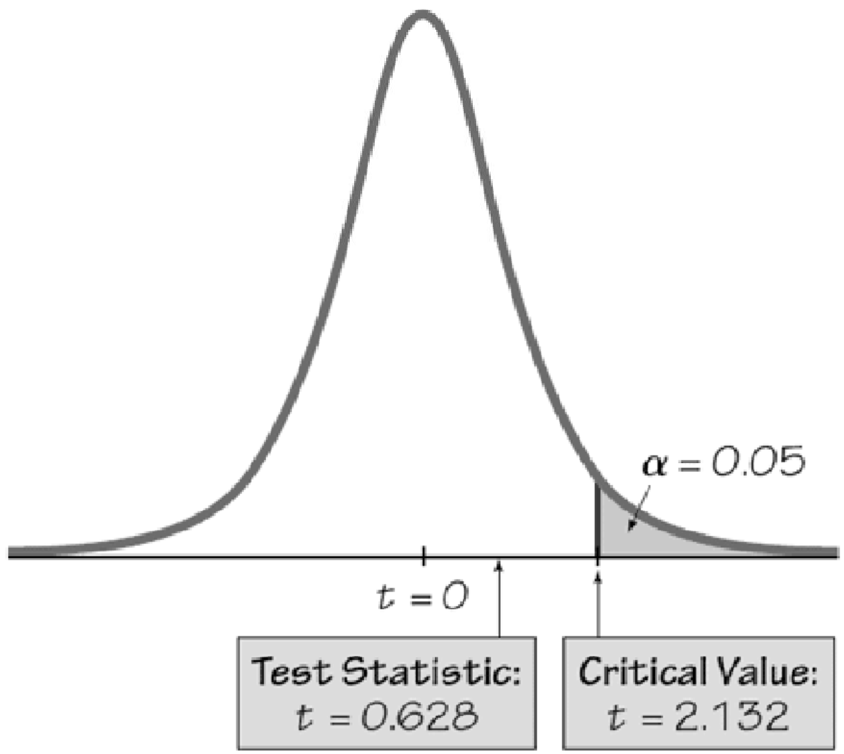
Step 6: Determine the value of the test statistic:

$$t = \frac{\bar{d} - \mu_d}{\dfrac{s_d}{\sqrt{n}}} = \frac{3.2 - 0}{\dfrac{11.4}{\sqrt{5}}} = 0.628$$

with df = 5 – 1 = 4

Example

Step 6: Using technology, the *P*-value is 0.282. Using the critical value method:



Example

Step 7: Because the *P*-value exceeds 0.05, or because the test statistic does not fall in the critical region, we fail to reject the null hypothesis.

Conclusion: There is not sufficient evidence to support the claim that for the population of heights of presidents and their main opponent, the differences have a mean greater than 0 cm.

In other words, presidents do not appear to be taller than their opponents.

Example

Confidence Interval: Support the conclusions with a 90% confidence interval estimate for μ_d.

$$E = t_{\alpha/2}\frac{s_d}{\sqrt{n}} = 2.132 \cdot \frac{11.4}{\sqrt{5}} = 10.8694$$

$$\bar{d} - E < \mu_d < \bar{d} + E$$

$$3.2 - 10.8694 < \mu_d < 3.2 + 10.8694$$

$$-7.7 < \mu_d < 14.1$$

Example

We have 90% confidence that the limits of –7.7 cm and 14.1 cm contain the true value of the difference in height (president's height – opponent's height).

See that the interval does contain the value of 0 cm, so it is very possible that the mean of the differences is equal to 0 cm, indicating that there is no significant difference between the heights.